

Efficient learning of ground and thermal states within phases of matter

Received: 16 October 2023

Accepted: 7 August 2024

Published online: 05 September 2024

Cambyse Rouzé¹✉, Daniel Stilck França²✉, Emilio Onorati³✉ & James D. Watson⁴✉

We consider two related tasks: (a) estimating a parameterisation of a given Gibbs state and expectation values of Lipschitz observables on this state; (b) learning the expectation values of local observables within a thermal or quantum phase of matter. In both cases, we present sample-efficient ways to learn these properties to high precision. For the first task, we develop techniques to learn parameterisations of classes of systems, including quantum Gibbs states for classes of non-commuting Hamiltonians. We then give methods to sample-efficiently infer expectation values of extensive properties of the state, including quasi-local observables and entropies. For the second task, we exploit the locality of Hamiltonians to show that M local observables can be learned with probability $1 - \delta$ and precision ε using $N = \mathcal{O}\left(\log\left(\frac{M}{\delta}\right)e^{\text{polylog}(\varepsilon^{-1})}\right)$ samples – exponentially improving previous bounds. Our results apply to both families of ground states of Hamiltonians displaying local topological quantum order, and thermal phases of matter with exponentially decaying correlations.

Tomography of quantum states is among the most important tasks in quantum information science. In quantum tomography, we have access to one or more copies of a quantum state and wish to understand the structure of the state. However, for a general quantum state, all tomographic methods inevitably require resources that scale exponentially in the size of the system^{1,2}. This is due to the curse of dimensionality: the number of parameters needed to fully describe a quantum system scales exponentially with the number of its constituent particles. Obtaining these parameters often necessitates the preparation and destructive measurement of exponentially many copies of the quantum system, as well as their storage in a classical memory. In particular, as the size of quantum devices continues to increase beyond what can be easily simulated classically, the community faces new challenges to characterise their output states in a robust and efficient manner.

Thankfully, only a few physically relevant observables are often needed to describe the physics of a system, e.g. its entanglement or energy. Recently, new methods of tomography have been proposed

which precisely leverage this important simplification to develop efficient state learning algorithms. One highly relevant development in this direction is that of classical shadows³. This new set of protocols allows for estimating physical observables of quantum spin systems that only depend on local properties from a number of measurements that scales logarithmically with the total number of qubits. However, the number of required measurements still faces an exponential growth with respect to the size of the observables that we want to estimate. Thus, using such protocols to learn the expectation values of physical observables that depend on more than a few qubits quickly becomes unfeasible.

For the task of state tomography, some simplification can be achieved from the fact that physically relevant quantum states, such as ground and Gibbs states of a locally interacting spin system, are themselves often described by a number of parameters which scales only polynomially with the number of qubits. From this observation, another direction in the characterisation of large quantum systems that has received considerable attention is that of Hamiltonian

¹Inria, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France. ²Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342 Lyon Cedex 07, France.

³Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany. ⁴University of Maryland, College Park QuICS 3353 Atlantic Building MD 20742-2420, USA. ✉e-mail: cambyse.rouze@inria.fr; daniel.stilck_franca@ens-lyon.fr; emilio.onorati@tum.de; jdwatson@umd.edu

learning and many-body tomography, where it was recently shown that it is possible to robustly characterise the interactions of a Gibbs state with a few samples^{4,5}. However, even for many-body states, recovery in terms of the trace distance requires a number of samples that scales polynomially in the number of qubits, in contrast to shadows for which the scaling is logarithmic.

These considerations naturally lead to the question of identifying settings where it is possible to combine the strengths of shadows and many-body tomography. In⁶, the authors proposed a first solution by combining these with new insights from the emerging field of quantum optimal transport. They obtained a tomography algorithm that only requires a number of samples that scales logarithmically in the system's size and learns all quasi-local properties of a state. These properties are characterised by so-called “Lipschitz observables”. However, that first step was confined to topologically trivial states such as high-temperature Gibbs states of commuting Hamiltonians or outputs of shallow circuits.

We also improve on the limitations of the tomography techniques discussed: we significantly extend these results beyond topologically trivial states to all states exhibiting exponential decay of correlations and the approximate Markov property. This result permits to significantly enlarge the class of states for which we know how to learn all quasi-local properties with a number of samples that scales polylogarithmically with the system's size. In particular, our results now also hold for classes of Gibbs states of non-commuting Hamiltonians. It also allows us to learn states with exponentially fewer samples in the quantum Wasserstein distance than was previously known before.

Beyond the limits of the specific tomography techniques sketched above, tomographical techniques by themselves are somewhat limited in that they tell us nothing about nearby related states - often states belong to a phase of matter in which the properties of the states vary smoothly and are in some sense “well behaved”, and we wish to learn properties of this entire phase of matter. A recent line of research in this direction that has gained significant attention from the quantum community is that of combining machine learning methods with the ability to sample complex quantum states from a phase of matter to efficiently characterise the entire phase^{7,8}, as well as using ML techniques to improve identifying phases of matter^{9–11} and approximating quantum states^{12–15}. It is well known that these tasks are computationally intractable in general^{16–18}, and so having access to data from an externally generated source could conceivably speed up these computations. A landmark result in this direction is¹⁹. There the authors showed how to use machine learning methods combined with classical shadows to learn local linear and nonlinear functions of states belonging to a gapped phase of matter with a number of samples that only grows logarithmically with the system's size. That is, given states from that phase drawn from a distribution and the corresponding parameters of the Hamiltonian, one can train a classical algorithm that would predict local properties of other points of the phase. However, there are some caveats to this scheme: (i) the scaling of the number of samples in terms of the precision is exponential, (ii) it does not immediately apply to phases of matter beyond gapped ground states, (iii) the results only come with guarantees on the errors in the prediction in expectation. That is, given another state sampled from the same distribution as the one used to train, only on average is the error made by the ML algorithm proven to be small.

In this work, we address all of these shortcomings. First, our result extends to thermal phases of matter which exhibit exponential decay of correlations, which includes thermal systems away from criticality/poles in the partition function^{20, Section 51}. Our result also extends to phases that satisfy a generalised version local topological quantum order^{21–23} which includes not only Gibbs states with exponentially decaying correlations, but also ground states of gapped Hamiltonians. Furthermore, the sample complexity of our algorithm is quasi-polynomial in the desired precision, which is an exponential

improvement over previous work¹⁹. Importantly, it also comes with point-wise guarantees on the quality of the recovery, as opposed to average guarantees.

Results

In this paper, we consider a quantum system defined over a D -dimensional finite regular lattice $\Lambda = [-L, L]^D$, where $n = (2L + 1)^D$ denotes the total number of qubits constituting the system. We assume for simplicity that each site of the lattice hosts a qubit, so that the total system's Hilbert space is $\mathcal{H}_\Lambda = \bigotimes_{j \in \Lambda} \mathbb{C}^2$, although all of the results presented here easily extend to qudits.

We prove two sets of results. In subsections ‘Preliminaries’, ‘Tomography: Optimal Tomography of Many-Body Quantum States’, and ‘Tomography: Beyond linear functionals’ we summarise our results on tomography; in particular, how to estimate all quasi-local properties of a given state given identical copies of it. This is the traditional setting of quantum tomography. We give our second set of results in subsections ‘Learning Phases: Learning Expectation Values of Parameterised Families of Many-Body Quantum Systems’ and ‘Learning Phases: Learning Beyond Exponentially Decaying Phases’ where we summarise our results on how to learn local properties of a class of states given samples from different states from that class. This is the setting of¹⁹ where ground states of gapped quantum phases of matter were studied. Here we consider (a) thermal phases of matter with exponentially decaying correlations and (b) ground states satisfying what we call ‘Generalised Approximate Local Indistinguishability’ (GALI). We show that ground state phases of gapped Hamiltonians satisfy this GALI condition.

Our focus in this work are nontrivial statements about what can be learned about many-body states of n qubits in the setting where we are only given $\Theta(\text{polylog}(n))$ copies. The common theme is that we will assume exponential decay of correlations for our class of states. This underlying assumption to our results is not an artificial constraint, but naturally appears in a broad range of physical settings, e.g., this is well-known to hold for ground states of uniformly gapped local Hamiltonians²⁴. Exponential decay of correlations also arises for high-temperature systems^{20,25–28}, 1D Gibbs states at any constant temperature^{20,29}, the Ising model away from critical points³⁰ and in the context of stationary states of Liouvillians that feature rapid mixing³¹.

Preliminaries

We first consider the task of obtaining a good approximation of expected values of extensive properties of a fixed unknown n -qubit state over Λ . The state is assumed to be a Gibbs state of an unknown local Hamiltonian $H(x) = \sum_{j \in \Lambda} h_j(x^{(j)})$, $x = \{x^{(j)}\} \in [-1, 1]^m$, defined through geometrically local interactions $h_j(x^{(j)})$, each depending on parameters $x^{(j)} \in [-1, 1]^\ell$ for some fixed integer ℓ and supported on a ball A_j around site $j \in \Lambda$ of radius r_0 . We also assume that the matrix-valued functions $x^{(j)} \mapsto h_j(x^{(j)})$ as well as their derivatives are uniformly bounded: $\|h_j\|_\infty, \|\nabla h_j\|_\infty \leq h$. Here, “uniformly” means that the bound applies to all system sizes, with constants that are independent of system size. The corresponding Gibbs state at inverse temperature $\beta > 0$, and the ground state as $\beta \rightarrow \infty$ take the form

$$\sigma(\beta, x) := \frac{e^{-\beta H(x)}}{\text{tr}[e^{-\beta H(x)}]} \quad \text{and} \quad \psi_g(x) := \lim_{\beta \rightarrow \infty} \sigma(\beta, x). \quad (1)$$

In the case when $[h_j(x_j), h_{j'}(x_{j'})] = 0$ for all $j, j' \in \Lambda$, the Hamiltonian $H(x)$ and its associated Gibbs states $\sigma(\beta, x)$ are said to be commuting.

We will often require that the states satisfy the property of exponential decay of correlations: for any two observables X_A , resp. X_B , supported on region A , resp. B ,

$$\text{Cov}_{\sigma(\beta, x)}(X_A, X_B) \leq C \min\{|A|, |B|\} \|X_A\|_\infty \|X_B\|_\infty e^{-\nu \text{dist}(A, B)}, \quad (2)$$

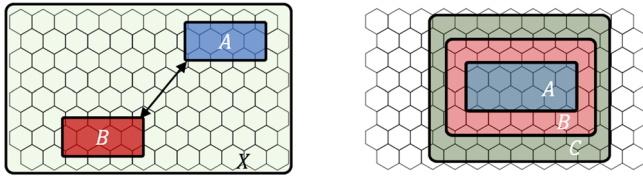


Fig. 1 | Illustration of the clustering of correlations and Markov Condition. Left: Exponential decay of correlations means the covariance in Eq. (2) must decrease exponentially between regions (A, B). Right: Roughly, the Markov Condition implies that the mutual information between regions (A, C), which are separated by a region (B), decreases sufficiently fast as the width of (B) increases.

for some constants $C, \nu > 0$, where $\text{dist}(A, B)$ denotes the distance between regions A and B , and where the covariance is defined by

$$\text{Cov}_\sigma(X, Y) := \frac{1}{2} \text{tr}[\sigma\{X - \text{tr}[\sigma X], Y - \text{tr}[\sigma Y]\}]. \quad (3)$$

We can also define a slightly stronger version of exponential decay:

Definition 1. (Uniform clustering) The Gibbs state $\sigma(\beta, x)$ is said to be uniformly $\zeta(\ell)$ -clustering if for any $X \subset A$ and any $A \subset X$ and $B \subset X$ such that $\text{dist}(A, B) \geq \ell$,

$$\text{Cov}_{\sigma(\beta, x)}(X_A, X_B) \leq \|X_A\|_\infty \|X_B\|_\infty \zeta(\ell) \quad (4)$$

for any X_A supported on A and X_B supported on B ($\sigma(\beta, x, X)$).

See Fig. 1 for an illustration. We note this condition implies eq. (2). As pointed out in ref. 32, this property is called uniform clustering to contrast with regular clustering property that usually only refers to properties of the state $\sigma(\beta, x)$.

Definition 2. (Uniform Markov condition) The Gibbs state $\sigma(\beta, x)$ is said to satisfy the uniform $\delta(\ell)$ -Markov condition if for any $ABC = X \subset A$ with B shielding A away from C (see Fig. 1) and such that $\text{dist}(i, j) \geq \ell$ for any $i \in A$ and $j \in C$, we have

$$I(A : C|B)_{\sigma(\beta, x)} \leq \delta(\ell). \quad (5)$$

We discuss these conditions more in Supplementary Note 3, III D3.

Part of the work will be focused on obtaining accurate estimates of quasi-local, extensive properties of quantum states, as such properties usually suffice to characterise physical properties of a state. For instance, we are interested in observables O that can be written as

$$O = \sum_{j \in \Lambda} O_j, \quad (6)$$

where O_j is supported on a ball B_j of radius $r_0 = \mathcal{O}(1)$ around site j and $\|O_j\|_\infty \leq 1$ for all $j \in \Lambda$. Such observables capture physical properties such as the energy, magnetisation, or the total number of particles. We might also consider nonlinear properties of the state, such as the average 2-Rényi entropy of a state ρ over regions:

$$\sum_{j \in \Lambda} -\log(\text{tr}[\rho_{B_j}^2]), \quad (7)$$

Such properties can be estimated from the reduced density matrices on each region B_j , but sometimes it is also important to consider quasi-local properties of quantum states. These are usually encoded in observables O like those of Eq. (6), but where each O_j is not necessarily strictly supported on B_j , but rather have a vanishing tail outside B_j . Such observables can arise if we e.g. evolve a strictly local observable by a Hamiltonian with algebraically decaying interactions.

The expectation value of such observables is then no longer determined by the reduced density of small regions.

Note, however, that when we consider learning protocols that have guarantees in terms of the trace distance, we have guarantees for all observables, not only for more physically motivated ones discussed above. Crucially for our results, extensive physical properties of a state are well-captured by the recently introduced class of Lipschitz observables^{33,34}.

Definition 3. (Lipschitz Observable³⁴) An observable L on \mathcal{H}_Λ is said to be *Lipschitz* if $\|L\|_{\text{Lip}} := \max_{i \in \Lambda} \min_{L_i} 2\|L - L_i \otimes I_i\|_\infty = \mathcal{O}(1)$, where i is the complement of the site i in Λ and the scaling is in terms of the number of qubits in the system.

In words, $\|L\|_{\text{Lip}}$ quantifies the amount by which the expectation value of L changes for states that are equal when tracing out one site. Lipschitz observables encompass a variety of physically motivated observables; indeed, for most physically motivated extensive quantities, it is not the case that flipping one qubit should substantially change the value of the expectation value. For instance, if we consider the energy with respect to a local Hamiltonian on a regular lattice, then acting with a unitary on one of the qubits can only change the energy by a fixed amount. This is in contrast with arbitrary observables, where one local change can lead to drastic changes in the expectation value. Thus, in some sense, Lipschitz observables capture the notion that for well-defined physical quantities, a local operation should only have a limited effect. One can readily see that for observables like that of Eq. (6), $\|L\|_{\text{Lip}} = \mathcal{O}(1)$. By a simple triangle inequality together with³⁴, Proposition 15, one can easily see that $\|L\|_\infty \leq n\|L\|_{\text{Lip}}$. Given the definition of the Lipschitz constant, we can also define the quantum Wasserstein distance of order 1 by duality³⁴.

Definition 4. (Quantum Wasserstein Distance³⁴) The *Quantum Wasserstein distance* between two n qubit quantum states ρ_0, ρ_1 is defined as $W_1(\rho_0, \rho_1) := \sup_{\|L\|_{\text{Lip}} \leq 1} \text{tr}[L(\rho_0 - \rho_1)]$. It satisfies $W_1(\rho_0, \rho_1) \leq n\|\rho_0 - \rho_1\|_1$.

Having

$$W_1(\rho, \sigma) = \mathcal{O}(\epsilon n) \quad (8)$$

is sufficient to guarantee that the expectation value of ρ and σ is the same up to an error that scales with the Lipschitz norm. This often implies a multiplicative error for *extensive, quasi-local* observables. Indeed, for any extensive, quasi-local observable O of the form of Eq. (6), we have that $\|O\|_{\text{Lip}} = \mathcal{O}(1)$, and thus from an inequality like that of Eq. (8) we have:

$$\text{tr}[O(\rho - \sigma)] \leq \|O\|_{\text{Lip}} W_1(\rho, \sigma) = \mathcal{O}(\epsilon n). \quad (9)$$

If we normalize the observable O to be traceless and have operator norm $\sim n$, then the expectation value of physical quantities will typically scale like $\sim n$, which justifies our claim that this recovery form will often lead to a multiplicative error.

This fact justifies why we focus on learning states up to an error $\mathcal{O}(\epsilon n)$ in Wasserstein distance instead of the usual trace distance bound of order $\mathcal{O}(\epsilon)$: although a trace distance guarantee of order $\mathcal{O}(\epsilon)$ would give the same error estimate, it requires exponentially more samples even for product states, as shown in ref. 6, Appendix G. We refer to Supplementary Note 2, for a more thorough discussion of why Lipschitz observables capture well quasi-local properties of quantum states.

Tomography: optimal tomography of many-body quantum states

We turn our attention to the problem of obtaining approximations to an unknown Gibbs state with as few samples as possible. That is, given copies of the unknown Gibbs state $\sigma(\beta, x)$, we want to learn a set of parameters x' such that the state $\sigma(\beta, x')$ provides a good estimate. In particular, ensuring that our approximation $\sigma(\beta, x')$ is $\mathcal{O}(n\epsilon)$ -close to $\sigma(\beta, x)$ in quantum Wasserstein distance guarantees that the description returned by our algorithm, in terms of parameters x' , satisfies $|f_L(\beta, x') - f_L(\beta, x)| = \mathcal{O}(n\epsilon)$ for linear functions of the form $f_L(\beta, x) := \text{tr}[L\sigma(\beta, x)]$ where L is a Lipschitz observable. For this task we only use measurements and classical post-processing on the copies of $\sigma(\beta, x)$ provided.

Our first main result is a method to learn Gibbs states with few copies of the unknown state:

Theorem 1. (Tomography algorithm for decaying Gibbs states (informal of Supplementary Theorem III.1)) For any unknown commuting Gibbs state $\sigma(\beta, x)$ satisfying exponential decay of correlations in Eq. (2), there exists an algorithm that outputs parameters x' such that the state $\sigma(\beta, x')$ approximates $\sigma(\beta, x)$ to precision $n\epsilon$ in Wasserstein distance with probability $1 - \delta$ with access to $N = \mathcal{O}(\log(\delta^{-1})\text{polylog}(n)\epsilon^{-2})$ samples of the state (see Supplementary Note 3, III D1). The result extends to non-commuting Hamiltonians whenever one of the following two assumptions is satisfied:

- the high-temperature regime, $\beta < \beta_c$ for some constant temperature $\beta_c = \mathcal{O}(1)$ (see Supplementary Note 3, III D2).
- uniform clustering of correlations (which implies Eq. (2)) & Markov conditions. In case (ii), we find good approximation guarantees under the following slightly worse scaling in the precision ϵ : $N = \mathcal{O}(\epsilon^{-4}\text{polylog}(n\delta^{-1}))$.

Proof ideas. The results for commuting Hamiltonians and in the high-temperature regime proceed directly from the following continuity bound on the Wasserstein distance between two arbitrary Gibbs states, whose proof requires the notion of quantum belief propagation in the non-commuting case (see Supplementary Corollary III.4): for any $x, y \in [-1, 1]^m$,

$$W_1(\sigma(\beta, x), \sigma(\beta, y)) = \beta \|x - y\|_{\ell_1} \mathcal{O}(\text{polylog}(n)). \quad (10)$$

Furthermore, this inequality is tight up to a $\text{polylog}(n)$ factor for $\beta = \mathcal{O}(1)$. Equation (10) reduces the problem of recovery in Wasserstein distance to that of recovering the parameters x up to an error $n\epsilon/\text{polylog}(n)$ in ℓ_1 distance. This is a variation of the Hamiltonian learning problem for Gibbs states^{5,35} which relies on lower bounding the ℓ_2 strong convexity constant for the log-partition function.

In⁴, the authors give an algorithm estimating the Hamiltonian parameters x with $e^{\mathcal{O}(\beta k^2)} \mathcal{O}(\beta^{-1} \log(\delta^{-1})n\epsilon^{-2})$ copies of $\sigma(\beta, x)$ up to ϵ in ℓ_∞ distance when $\sigma(\beta, x)$ belongs to a family of commuting, k -local Hamiltonians on a D -dimensional lattice. If we assume $m = \mathcal{O}(n)$, this translates to an algorithm with sample complexity $e^{\mathcal{O}(\beta k^2)} \mathcal{O}(\epsilon^{-2} \text{polylog}(\delta^{-1})n)$ to learn x up to $n\epsilon$ in ℓ_1 distance. It should also be noted that the time complexity of the algorithm in ref. 4 is $\mathcal{O}(ne^{\mathcal{O}(\beta k^2)} \epsilon^{-2} \text{polylog}(\delta^{-1})n)$. Thus, any commuting model at constant temperature satisfying exponential decay of correlations can be efficiently learned with $\text{polylog}(n)$ samples. We refer the reader to Supplementary Note 5 for more information and classes of commuting states that satisfy exponential decay of correlations. In the high-temperature regime, we rely on a result of⁵ where the authors give a computationally efficient algorithm to learn x up to error ϵ in ℓ_∞ norm from $\mathcal{O}(\epsilon^{-2} \text{polylog}(\delta^{-1})n)$ samples. This again translates to a $\mathcal{O}(n)$ error in ℓ_1 norm thanks to (10).

Furthermore, in Supplementary Note 3, III D3 we more directly extend the strategy of³⁵ by introducing the notion of a W_1 strong

convexity constant for the log-partition function and showing that it scales linearly with the system size under (a) uniform clustering of correlations and (b) uniform Markov condition. This result also generalises the strategy of⁶ which relied on the existence of a so-called transportation cost inequality previously shown to be satisfied for commuting models at high-temperature. For the larger class of states satisfying conditions (a) and (b), we are able to find x' s.t. $W_1(\sigma(\beta, x), \sigma(\beta, x')) = \mathcal{O}(n\epsilon)$ with $\mathcal{O}(\epsilon^{-4} \text{polylog}(\delta^{-1})n)$ samples. Note that the uniform Markov condition is expected to hold for a large class of models that goes beyond high-temperature Gibbs states^{25,36}. We believe that our result is also of interest for classical models. The last years have seen a flurry of results on learning classical Gibbs states under various metrics, with a particular focus on learning the parameters in some ℓ_p norm^{37–41}. But, to the best of our knowledge, the learning in W_1 was not considered, particularly the quantum version of the Wasserstein distance. Furthermore, there are phases of classical Ising models that exhibit exponential decay of correlations but no polynomial-time algorithms to sample from the underlying Gibbs states are known^{42,43}. Note, however, that here we are assuming that we have access to samples from the distribution, which makes it possible to devise efficient learning algorithms even when sampling is expected to be hard³⁷. Thus, generally speaking, there is no simple connection between the efficient simulation (classical or quantum) of a class of Gibbs states and the existence of efficient Hamiltonian learning algorithms.

Tomography: Beyond linear functionals

So far, we considered properties of the quantum system which could be related to local linear functionals of the unknown state. In^{3,19}, the authors propose a simple trick in order to learn non-linear functionals of many-body quantum systems, e.g. their entropy over a small sub-region. However, such methods require a number of samples scaling exponentially with the size of the subregion, and thus very quickly become inefficient as the size of the region increases. Here instead, we use³⁴, Theorem 1, where the authors show the continuity bound on the von Neumann entropy S :

$$|S(\rho) - S(\sigma)| \leq g(W_1(\rho, \sigma)) + W_1(\rho, \sigma) \log(4n), \quad (11)$$

where $g(t) = (t+1)\log(t+1) - t\log(t)$, together with the following Wasserstein continuity bound in order to estimate the entropic quantities of Gibbs states over regions of arbitrary size (see Supplementary Corollary III.6): assuming Eq. (2), for any region R of the lattice and any two $x, y \in [-1, 1]^m$

$$W_1(\text{tr}_{R^c}(\sigma(\beta, x)), \text{tr}_{R^c}(\sigma(\beta, y))) \leq \|x|_{\mathcal{R}(r_R)} - y|_{\mathcal{R}(r_R)}\|_{\ell_1} \text{polylog}(|R(r_R)|), \quad (12)$$

where $r_R = \max\{r_0, 2\xi \log(2|R|C_1\|x|_{\mathcal{R}(r_0)} - y|_{\mathcal{R}(r_0)}\|_{\ell_1}^{-1})\}$ with r_0 being the smallest integer such that $x|_{\mathcal{R}(r_0)} \neq y|_{\mathcal{R}(r_0)}$, and where $\mathcal{R}(r_R) := \{x^{(j)} | \text{supp}(h_j(x^{(j)})) \cap R(r_R) \neq \emptyset\}$ is the set of parameters $x^{(j)}$ describing the Hamiltonian in the region $R(r_R)$, and where $R(r_R) := \{i \in A : \text{dist}(i, R) \leq r_R\}$. $C_1, \xi > 0$ are $\mathcal{O}(1)$ constants.

Let us recall a few definitions: denoting by $\rho_R := \text{tr}_{R^c}(\rho)$ the marginal of a state $\rho \in \mathcal{D}(\mathcal{H}_A)$ on a region $R \subset A$, and given separated regions $A, B, C \subset A$ of the lattice: $S(A)_\rho := -\text{tr}[\rho_A \log \rho_A]$ is the von Neumann entropy of ρ on A , $S(A|B)_\rho := S(AB)_\rho - S(B)_\rho$ is the conditional entropy on region A conditioned on region B , $I(A: B)_\rho := S(A)_\rho + S(B)_\rho - S(AB)_\rho$ is the mutual information between regions A and B , and $I(A: B|C)_\rho := S(AC)_\rho + S(BC)_\rho - S(C)_\rho - S(ABC)_\rho$ is the conditional mutual information between regions A and B conditioned on region C . Equation (11) can be combined together with Eq. (12) to get the following:

Corollary 1. Assume the decay of correlations holds uniformly, as specified in Eq. (2), for all $\{\sigma(\beta, x)\}_{x \in [-1, 1]^m}$, $m = \mathcal{O}(n)$. Then, in the

notations of the above paragraph, for any two Gibbs states $\sigma(\beta, x)$ and $\sigma(\beta, y)$, $x, y \in [-1, 1]^m$, and any region $A \subset \Lambda$:

$$|S(A)_{\sigma(\beta, x)} - S(A)_{\sigma(\beta, y)}| = \|x|_{\mathcal{R}(R_R)} - y|_{\mathcal{R}(R_R)}\|_{\ell_1} \mathcal{O}(\text{polylog}(|\mathcal{R}(R_R)|)), \quad (13)$$

for $R \equiv A$. The same conclusion holds for $|S(A|B)_{\sigma(\beta, x)} - S(A|B)_{\sigma(\beta, y)}|$, $(R \equiv AB)$, $|I(A: B)_{\sigma(\beta, x)} - I(A: B)_{\sigma(\beta, y)}|$ ($R \equiv AB$), and $|I(A: B|C)_{\sigma(\beta, x)} - I(A: B|C)_{\sigma(\beta, y)}|$ ($R \equiv ABC$).

Thus, given an estimate y of x satisfying $\|x - y\|_{\ell_\infty} = \mathcal{O}(\varepsilon/\text{polylog}(n))$, we can also approximate entropic quantities of the Gibbs state to a multiplicative error. Furthermore, we remark that for many classes of Gibbs states, including high-temperature and commuting models^{4–6}, we can obtain such an estimate with $\mathcal{O}(\varepsilon^{-2} \text{polylog}(n))$ samples, and it remains an important open problem to establish such a result in general for quantum models. More generally, entropic continuity bounds can be directly used together with Theorem 1(ii) in order to estimate entropic properties of Gibbs states satisfying both uniform clustering of correlations and the approximate Markov condition (see Supplementary Note 3, III D 3 for details).

Learning phases: learning expectation values of parameterised families of many-body quantum systems

Next, we turn our attention to the task of learning Gibbs or ground states of a parameterised Hamiltonian $H(x)$ known to the learner and sampled according to the uniform distribution U over some region $x \in \Phi := \prod_{i=1}^m [-1 + x_i^0, 1 + x_i^0]$ (where here $\prod_{i=1}^m$ represents the Cartesian product over sets). More general distributions can also be dealt with under a condition of anti-concentration, see Supplementary Note 4. Here we restrict our results to local observables of the form $O = \sum_{i=1}^M O_i$ where $R_i := \text{supp}(O_i)$ is contained in a ball of diameter independent of the system size. The setup in this section is similar to¹⁹. The idea is that we have access to some samples of a state chosen from different values of the parameterised Hamiltonian, and we want to use these to learn observables everywhere in the parameter space with high precision. We then want to know: what is the minimum number of samples drawn from this distribution which allows us to accurately predict expectation values of local observables for every choice of parameters?

The learner is given samples $\{(x_i, \sigma(\beta, x_i))\}_{i=1}^N$, where the parameters $x_i \sim U$, and their task is to learn $f_O(x) := \text{tr}[\sigma(\beta, x)O]$ for an arbitrary value of $x \in \Phi$ and an arbitrary local observable O . We assume that everywhere in the parameter space $x \in \Phi$ the Gibbs states are in the same phase of exponentially decaying correlations. Note that this does not necessarily imply the existence of a fully polynomial time approximation scheme, and finding under which conditions such algorithms exist is still a very active area of research⁴². As discussed at the beginning of the subsection ‘Tomography: Optimal Tomography of Many-Body Quantum States’, exponentially decaying correlations is a natural condition for many physical systems.

Theorem 2. (Learning algorithm for quantum Gibbs states informal version of Supplementary Theorem IV.4) With the conditions of the previous paragraph, given a set of N samples $\{(x_i, \sigma(\beta, x_i))\}_{i=1}^N$, where $\sigma(\beta, x_i)$ can be stored efficiently classically, and $N = \mathcal{O}(\log(\frac{M}{\varepsilon}) \log(\frac{n}{\varepsilon}) e^{\text{polylog}(\varepsilon^{-1})})$, there exists an algorithm that, on input $x \in \Phi$ and a local observable $O = \sum_{i=1}^M O_i$, produces an estimator \hat{f}_O such that, with probability $(1 - \delta)$,

$$\sup_{x \in \Phi} |f_O(x) - \hat{f}_O(x)| \leq \varepsilon \sum_{i=1}^M \|O_i\|_\infty. \quad (14)$$

Moreover, the samples $\sigma(\beta, x_i)$ are efficiently generated from measurements of the Gibbs states $\{\sigma(\beta, x_i)\}_{i=1}^N$ followed by classical post-processing.

Proof ideas. Our estimator \hat{f}_O is constructed as follows: during a training stage, we pick N points $Y_1, \dots, Y_N \sim U$ and estimate the reduced Gibbs states over large enough enlargements $R_i \partial$ of the supports $\mathcal{R}_i := \{x^{(j)} | \text{supp}(h_j(x^{(j)})) \cap R_i \neq \emptyset\} \cap [x - \varepsilon, x + \varepsilon]^m$ of the observables O_i . Due to the anti-concentration property of the uniform distribution, the probability that a small region $\mathcal{R}_i \partial$ in parameter space contains t variables Y_{i_1}, \dots, Y_{i_t} becomes large for $N \approx \log(M)$.

We then run a robust version of the classical shadow tomography protocol on those states in order to construct efficiently describable and computable product matrices $\tilde{\sigma}(\beta, Y_{i_1}), \dots, \tilde{\sigma}(\beta, Y_{i_t})$. This is devised in Supplementary Proposition 4.3. Then for any region R_i , we select the shadows $\tilde{\sigma}(\beta, Y_{i_1}), \dots, \tilde{\sigma}(\beta, Y_{i_t})$ such that the restriction of each Y_{i_k} to just the parameters in R_i are close to that of the target state and construct the empirical average $\tilde{\sigma}_{R_i}(x) := \frac{1}{t} \sum_{j=1}^t \text{tr}_{R_i^c} [\tilde{\sigma}(\beta, Y_{i_j})]$. That is, the samples which are included in this sum satisfy that the difference $\|Y_{j|R_i} - x|_{R_i}\|_\infty$ is sufficiently small. Using belief propagation methods (see Supplementary Proposition IV.3), it is possible to show that exponential decay of correlations ensures that the estimator is a good approximation to local observables. Thus such operators can be well approximated using the reduced state $\text{tr}_{R_i^c} \sigma(\beta, x)$ for $t \approx \log(n)$. The estimator \hat{f}_O is then naturally chosen as $\hat{f}_O(x) := \sum_{i=1}^M \text{tr}[O_i \tilde{\sigma}_{R_i}(x)]$. A key part of the proof is demonstrating that exponential decay of correlations implies that $f_O(x)$ does not change too much as x varies. See Fig. 2 for an outline of the algorithm.

We note that the algorithm here strictly works when there is only one phase of matter. In the presence of more than one phase, the predictor will give less accurate results as it approaches the phase boundary, however, it the results of theorem 2 will hold within the bulk of each phase. Finally, we note that once we have access to the shadows, the algorithm simply performs a nearest-neighbour search with respect to the parameter of interest, followed by constructing the estimator, giving a runtime of at most $\mathcal{O}(nN)$.

Learning phases: learning beyond exponentially decaying phases

So far we have only discussed results for families of thermal states satisfying exponential decay of correlations. It would be desirable to extend our results to phases for which this is not generally known to hold, and to be able to extend our results to zero temperature systems. We introduce a new condition called *generalised approximate local indistinguishability* (GALI), under which learning local observables from samples can be done efficiently. GALI can be shown to hold for all gapped ground state phases of matter and all thermal states with exponentially decaying correlations. We refer the reader to Supplementary Note 5 for further details.

Definition 5. (Generalised approximate local indistinguishability (GALI)) For $x_1^0, \dots, x_m^0 \in \mathbb{R}$ let $\Phi := \prod_{i=1}^m [-1 + x_i^0, 1 + x_i^0]$ and for $x \in \Phi$ define $\rho(x)$ as either the ground state or thermal state of the local Hamiltonian $H(x)$. We say that the family of states $\rho(x)$ satisfies generalised approximate local indistinguishability (GALI) with decay function f if for any region R and $r \in \mathbb{N}$ there is a set of parameters $x_{\mathcal{R}(R)^c}^*$ s.t. for all O supported on R and $f_O(x) := \text{tr}[O \rho(x)]$ the following bound holds:

$$\sup_{x \in \Phi} |f_O(x) - f_O((x|_{\mathcal{R}(R)} x_{\mathcal{R}(R)^c}^*))| \leq |R| f(r) \|O\|_\infty, \quad (15)$$

for a function f s.t. $\lim_{r \rightarrow \infty} f(r) = 0$.

Under the GALI assumption, we are able to prove the following generalisation of theorem 2:

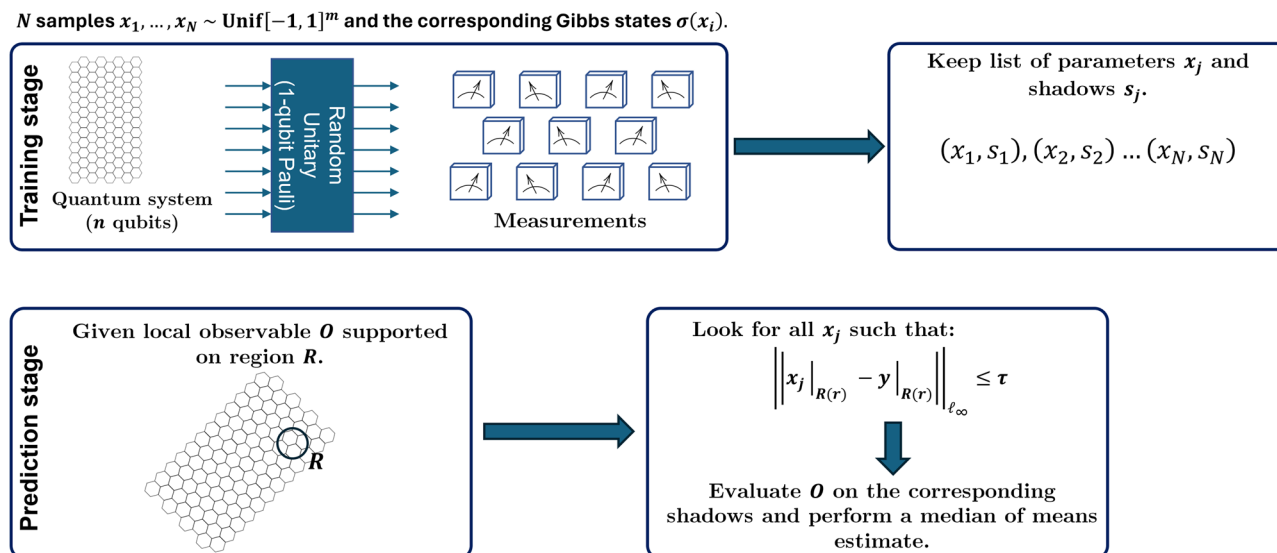


Fig. 2 | Schematic representation of our algorithm to learn phases of matter. The training stage just consists of collecting shadows corresponding to various parameters. In the prediction stage, given an observable O and corresponding parameter y supported on a region R , we search for parameters x^i we sampled that

have parameters close to y on an enlarged region, $R(r)$ around R and compute the expectation value of O on the corresponding shadows. The prediction is then a median of means estimate on the values. Note that no machine learning techniques are required for the estimate.

Theorem 3. Let $\rho(x)$, $x \in \Phi := \prod_{i=1}^m [-1+x_i^0, 1+x_i^0]$, be a family of ground states or Gibbs states satisfying GALI, as per definition 5 with $f(r) = \mathcal{O}(e^{-r/\xi})$ for some correlation length $\xi > 0$. Given a set of N samples $\{x_i, \tilde{\rho}(x_i)\}_{i=1}^N$, where $\tilde{\rho}(x_i)$ can be stored efficiently classically, and $N = \mathcal{O}\left(\log\left(\frac{M}{\delta}\right) \log\left(\frac{n}{\delta}\right) e^{\text{polylog}(e^{-1})}\right)$, there exists an algorithm that, on input $x \in \Phi$ and a local observable $O = \sum_{i=1}^M O_i$, produces an estimator \hat{f}_O such that, with probability $(1 - \delta)$,

$$\sup_{x \in [-1, 1]^m} |f_O(x) - \hat{f}_O(x)| \leq \varepsilon \sum_{i=1}^M \|O_i\|_\infty. \quad (16)$$

Moreover, the samples $\tilde{\rho}(x_i)$ are efficiently generated from measurements of the Gibbs states $\{\rho(x_i)\}_{i=1}^N$ followed by classical post-processing.

Since ground states of Hamiltonians with a non-zero spectral gap satisfy GALI, we can efficiently learn families of gapped ground states with the methods in this work. We refer the reader to Supplementary Note 5 B for proofs. We also note that, as per the thermal state learning algorithm, in the presence of multiple phases the predictions will be good in the bulk, but may break down approaching the boundary.

Discussion

In this paper we contributed to the tasks of tomography and learnability of quantum many-body states by combining previous techniques with approaches not considered so far in this field, in order to obtain novel and powerful features.

First, we extended the results of⁶ on the efficient tomography of high-temperature commuting Gibbs states to Gibbs states with exponentially decaying correlations. This result permits to significantly enlarge the class of states for which we know how to learn all quasi-local properties with a number of samples that scales polylogarithmically with the system's size. In particular, our results now also hold for classes of Gibbs states of non-commuting Hamiltonians. As we require exponentially fewer samples to learn in the Wasserstein metric when compared with the usual trace distance and still recover essentially all physically relevant quantities associated to the states, we

hope that our results motivate the community to consider various tomography problems in the Wasserstein instead of trace distance.

As we achieved this result by reducing the problem of learning the states to learning the parameters of the Hamiltonian in ℓ_1 , we hope our work further motivates the study of the Hamiltonian learning problem in ℓ_1 -norm with polylog samples. 1D Gibbs states are a natural place to start, but obtaining Hamiltonian learning algorithms just departing from exponential decay of correlations would provide us with a complete picture. In Supplementary Note 3, III D 3, we also partially decoupled the Hamiltonian learning problem from the W_1 learning one by resorting to the uniform Markov condition. Thus, it would be important to establish the latter for a larger number of systems.

It would be interesting to investigate the sharpness of our bounds, and to understand if exponential decay of correlations is really necessary. One way of settling this question would be to prove polynomial lower bounds for learning in Wasserstein distance for states at critical temperatures.

Second, we improved the results of Huang et al.¹⁹ for learning a class of states in several directions, including the sample-complexity scaling in precision, the classes of states it applies to and the form of the recovery guarantee. In particular, the results now apply to Gibbs states, which are the states of matter commonly encountered experimentally. Additionally, we achieve recovery guarantees in the worst case, whereas Huang et al.¹⁹ only gives guarantees on the average-case error. Interestingly, we did not need to resort to complex machine learning techniques to achieve an exponentially better scaling in precision by making arguably mild assumptions on the distributions the states are drawn from. Our algorithm is essentially a simple nearest-neighbour algorithm. Although the results proved here push the state-of-the-art of learning quantum states, we believe that our methods, for instance the novel continuity bounds for various local properties of quantum many-body states, will find applications in other areas of quantum information.

Beyond the GALI thermal phases and ground states studied here, it would be interesting to find other families of states which can be efficiently learned, and indeed if more restrictive assumptions on the parameterisation of Hamiltonians can result in more efficient learning. One interesting open problem that goes beyond the present paper's

scope is finding families of states satisfying GALI without belonging to a common gapped phase of matter. If such a family existed, it would clarify the differences between our framework and that of⁴⁴. We also refer to work on learning phases of matter, where phase here is defined under a rapid-mixing condition⁴⁵. In addition, it would be interesting to derive lower bounds on the sample complexity for the problem of learning a quantum phase as in this work. Finally, we realise that although the results proved here are for lattice systems, they almost certainly generalise to non-lattice configurations of particles.

We also recognise independent, concurrent work by Lewis et al.⁴⁴. Here the authors consider the same setup of gapped ground states as Huang et al.¹⁹ and also improved the sample complexity to achieve the same scaling as Theorem 3. However, their result is not directly comparable to ours. We emphasise that Lewis et al.⁴⁴ consider gapped, ground state phases, whereas our work includes thermal phases. We also note they remove all conditions on the prior distribution over the samples x , whereas we still need to assume a type of mild anti-concentration over the local marginals. However, their result is still stated as an $\|\cdot\|_{L^2}$ -bound, whereas our more straightforward approximation tools allow us to get stronger bounds in $\|\cdot\|_{\infty}$. Conceptually speaking, our methods for approximating local expectation values requires no knowledge of machine learning techniques. Our work also shows that it is possible to go beyond gapped quantum phases and learn thermal phases.

References

- Haah, J., Harrow, A. W., Ji, Z., Wu, X. & Yu, N. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory* **1**–1 (2017).
- O'Donnell, R. & Wright, J. Efficient quantum tomography. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, 899–912 (Association for Computing Machinery, New York, NY, USA <https://doi.org/10.1145/2897518.2897544> 2016).
- Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
- Anshu, A., Arunachalam, S., Kuwahara, T. & Soleimanifar, M. Efficient learning of commuting Hamiltonians on lattices. *Electronic notes* https://anuraganshu.seas.harvard.edu/files/anshu/files/learning_commuting_hamiltonian.pdf (2021).
- Haah, J., Kothari, R. & Tang, E. Optimal learning of quantum hamiltonians from high-temperature gibbs states. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, 135–146 (IEEE, 2022).
- Rouzé, C. & França, D. S. Learning quantum many-body systems from a few copies. *Quantum* **8**, 1319 (2024).
- Biamonte, J. et al. Quantum machine learning. *Nature* **549**, 195–202 (2017).
- Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431–434 (2017).
- Rodriguez-Nieva, J. F. & Scheurer, M. S. Identifying topological order through unsupervised machine learning. *Nat. Phys.* **15**, 790–795 (2019).
- Rem, B. S. et al. Identifying quantum phase transitions using artificial neural networks on experimental data. *Nat. Phys.* **15**, 917–920 (2019).
- Dong, X.-Y., Pollmann, F. & Zhang, X.-F. Machine learning of quantum phase transitions. *Phys. Rev. B* **99**, 121104 (2019).
- Gao, X. & Duan, L.-M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **8**, 662 (2017).
- Park, C.-Y. & Kastoryano, M. J. Geometry of learning neural quantum states. *Phys. Rev. Res.* **2**, 023232 (2020).
- Barr, A., Gispén, W. & Lamacraft, A. Quantum ground states from reinforcement learning. In *Mathematical and Scientific Machine Learning*, 635–653 (PMLR, 2020).
- Nomura, Y., Yoshioka, N. & Nori, F. Purifying deep boltzmann machines for thermal quantum states. *Phys. Rev. Lett.* **127**, 060601 (2021).
- Ambainis, A. On physical problems that are slightly more difficult than qma. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, 32–43 (IEEE, 2014).
- Watson, J. D. & Bausch, J. The complexity of approximating critical points of quantum phase transitions. *arXiv preprint arXiv:2105.13350* (2021).
- Bravyi, S., Chowdhury, A., Gosset, D. & Wocjan, P. Quantum hamiltonian complexity in thermal equilibrium. *Nat. Phys.* **18**, 1367–1370 (2022).
- Huang, H.-Y., Kueng, R., Torlai, G., Albert, V. V. & Preskill, J. Provably efficient machine learning for quantum many-body problems. *Science* **377** (2022).
- Harrow, A. W., Mehraban, S. & Soleimanifar, M. Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, 378–386 (2020).
- Michalakis, S. & Zwolak, J. P. Stability of frustration-free Hamiltonians. *Commun. Math. Phys.* **322**, 277–302 (2013).
- Bravyi, S., Hastings, M. B. & Michalakis, S. Topological quantum order: stability under local perturbations. *J. Math. Phys.* **51**, 093512 (2010).
- Nachtergaele, B., Sims, R. & Young, A. Quasi-locality bounds for quantum lattice systems. part ii. perturbations of frustration-free spin models with gapped ground states. In *Annales Henri Poincaré*, vol. 23, 393–511 (Springer, 2022).
- Hastings, M. B. & Koma, T. Spectral gap and exponential decay of correlations. *Commun. Math. Phys.* **265**, 781–804 (2006).
- Kuwahara, T., Kato, K. & Brandão, F. G. S. L. Clustering of conditional mutual information for quantum gibbs states above a threshold temperature. *Phys. Rev. Lett.* **124**, 220601 (2020).
- Kliesch, M., Gogolin, C., Kastoryano, M., Riera, A. & Eisert, J. Locality of temperature. *Phys. Rev. X* **4**, 031019 (2014).
- Migdal, A. Correlation functions in the theory of phase transitions: Violation of the scaling laws. *SOVIET PHYS. JETP* **32** (1971).
- Shao, S. & Sun, Y. Contraction: A unified perspective of correlation decay and zero-freeness of 2-spin systems. *J. Stat. Phys.* **185**, 1–25 (2021).
- Bluhm, A., Capel, Á. & Pérez-Hernández, A. Exponential decay of mutual information for Gibbs states of local Hamiltonians. *Quantum* **6**, 650 (2022).
- Duminil-Copin, H., Goswami, S. & Raoufi, A. Exponential decay of truncated correlations for the ising model in any dimension for all but the critical temperature. *Commun. Math. Phys.* **374**, 891–921 (2020).
- Kastoryano, M. J. & Eisert, J. Rapid mixing implies exponential decay of correlations. *J. Math. Phys.* **54**, 102201 (2013).
- Brandão, F. G. & Kastoryano, M. J. Finite correlation length implies efficient preparation of quantum thermal states. *Commun. Math. Phys.* **365**, 1–16 (2018).
- Rouzé, C. & Datta, N. Concentration of quantum states from quantum functional and transportation cost inequalities. *J. Math. Phys.* **60**, 012202 (2019).
- De Palma, G., Marvian, M., Trevisan, D. & Lloyd, S. The quantum Wasserstein distance of order 1. *IEEE Trans. Inf. Theory* **67**, 6627–6643 (2021).
- Anshu, A., Arunachalam, S., Kuwahara, T. & Soleimanifar, M. Sample-efficient learning of interacting quantum systems. *Nat. Phys.* **17**, 931–935 (2021).

36. Kato, K. & Brandao, F. G. Quantum approximate Markov chains are thermal. *Commun. Math. Phys.* **370**, 117–149 (2019).
37. Bresler, G. Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the forty-seventh annual ACM symposium on Theory of Computing* (ACM, <https://doi.org/10.1145/2746539.2746631> 2015).
38. Lokhov, A. Y., Vuffray, M., Misra, S. & Chertkov, M. Optimal structure and parameter learning of Ising models. *Sci. Adv.* **4** <https://doi.org/10.1126/sciadv.1700791> (2018).
39. Klivans, A. & Meka, R. Learning graphical models using multiplicative weights. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, 343–354 (IEEE, 2017).
40. Wu, S., Sanghavi, S. & Dimakis, A. G. Sparse logistic regression learns all discrete pairwise graphical models. *Advances in Neural Information Processing Systems* **32** (2019).
41. Dagan, Y., Daskalakis, C., Dikkala, N. & Kandiros, A. V. Learning Ising models from one or multiple samples (2020). <https://arxiv.org/abs/2004.09370>.
42. Helmuth, T., Perkins, W. & Regts, G. Algorithmic Pirogov-Sinai theory. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, 1009–1020 (Association for Computing Machinery, New York, NY, USA, <https://doi.org/10.1145/3313276.3316305> 2019).
43. Lubetzky, E., Martinelli, F., Sly, A. & Toninelli, F.-L. Quasi-polynomial mixing of the 2d stochastic Ising model with “plus” boundary up to criticality. *J. Eur. Math. Soc.* **15**, 339–386 (2013).
44. Lewis, L. et al. Improved machine learning algorithm for predicting ground state properties. *Nat. Commun.* **15**, 895 (2024).
45. Onorati, E., Rouzé, C., França, D. S. & Watson, J. D. Provably efficient learning of phases of matter via dissipative evolutions. *arXiv pre-print arXiv:2311.07506* (2023).

Acknowledgements

We gratefully recognise useful discussions with Hsin-Yuan (Robert) Huang, Haonan Zhang and Jens Eisert. We thank Laura Lewis, Viet T. Tran, Sebastian Lehner, Richard Kueng, Hsin-Yuan (Robert) Huang, and John Preskill for sharing a preliminary of the manuscript⁴⁴. C.R. acknowledges financial support from a Junior Researcher START Fellowship from the DFG cluster of excellence 2111 (Munich Centre for Quantum Science and Technology), from the ANR project QTraj (ANR-20-CE40-0024-01) of the French National Research Agency (ANR), as well as from the Humboldt Foundation. DSF is supported by France 2030 under the French National Research Agency award number “ANR-22-PNCQ-0002”. E.O. is supported by the Munich Quantum Valley and the Bavarian state government, with funds from the Hightech Agenda Bayern Plus, and by the European Research Council under grant agreement no. 101001976 (project EQUIPTNT). JDW acknowledges support from the United States Department of Energy, Office of Science,

Office of Advanced Scientific Computing Research, Accelerated Research in Quantum Computing programme, and also NSF QLCI grant OMA-2120757.

Author contributions

C.R. and J.D.W. were the primary contributors to this work. D.S.F. and E.O. contributed equally.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51439-x>.

Correspondence and requests for materials should be addressed to Cambyse Rouzé, Daniel Stilck França, Emilio Onorati or James D. Watson.

Peer review information *Nature Communications* thanks Yanming Che, Yu Tong and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024